Welfare Structure in Two-sided Random Matching Markets

Geng Zhao UC Berkeley

Joint work with Itai Ashlagi (Stanford) and Mark Braverman (Princeton)

From Matchings to Markets CIRM, December 2023

Ashlagi, Braverman, Zhao

Welfare Structure in Two-sided Random Matching Markets

1 / 20

- - E



• Applications: online dating, school choice, resident matching, etc.



- Applications: online dating, school choice, resident matching, etc.
- Stability: resistance to single-party deviation



- Applications: online dating, school choice, resident matching, etc.
- Stability: resistance to single-party deviation
- Often study random preferences: easier to generalize; "quality" vs "luck"



- Applications: online dating, school choice, resident matching, etc.
- Stability: resistance to single-party deviation
- Often study random preferences: easier to generalize; "quality" vs "luck"



- Applications: online dating, school choice, resident matching, etc.
- Stability: resistance to single-party deviation
- Often study random preferences: easier to generalize; "quality" vs "luck"



• *n* agents on each side ("women" and "men") with **logit-based** random preferences



- *n* agents on each side ("women" and "men") with **logit-based** random preferences
 - Woman i's preference is generated from a logit model with stochastic vector $\hat{\mathbf{a}}_i = (\hat{a}_{ij})_{j \in [n]} \in \mathbb{R}^n_+$



- *n* agents on each side ("women" and "men") with **logit-based** random preferences
 - Woman i's preference is generated from a logit model with stochastic vector $\hat{\mathbf{a}}_i = (\hat{a}_{ij})_{j \in [n]} \in \mathbb{R}^n_+$



- *n* agents on each side ("women" and "men") with **logit-based** random preferences
 - Woman *i*'s preference is generated from a logit model with stochastic vector $\hat{\mathbf{a}}_i = (\hat{a}_{ij})_{j \in [n]} \in \mathbb{R}^n_+$
 - E.g., $\hat{\bf a}_i=(1/2,1/3,1/6)\,\Rightarrow$ man 1 is 3 times more attractive to woman i (ex ante) than man 2



- *n* agents on each side ("women" and "men") with **logit-based** random preferences
 - Woman i's preference is generated from a logit model with stochastic vector $\hat{\mathbf{a}}_i = (\hat{a}_{ij})_{j \in [n]} \in \mathbb{R}^n_+$
 - E.g., $\hat{\bf a}_i=(1/2,1/3,1/6)$ \Rightarrow man 1 is 3 times more attractive to woman i (ex ante) than man 2
 - $\hat{\mathbf{A}}$: matrix with $\hat{\mathbf{a}}_i$'s as rows; similarly $\hat{\mathbf{B}}$ for men

A B M A B M



- *n* agents on each side ("women" and "men") with **logit-based** random preferences
 - Woman *i*'s preference is generated from a logit model with stochastic vector $\hat{\mathbf{a}}_i = (\hat{a}_{ij})_{j \in [n]} \in \mathbb{R}^n_+$
 - E.g., $\hat{\bf a}_i=(1/2,1/3,1/6)\,\Rightarrow$ man 1 is 3 times more attractive to woman i (ex ante) than man 2
 - $\hat{\mathbf{A}}$: matrix with $\hat{\mathbf{a}}_i$'s as rows; similarly $\hat{\mathbf{B}}$ for men
- Stability: no blocking pairs



- *n* agents on each side ("women" and "men") with **logit-based** random preferences
 - Woman *i*'s preference is generated from a logit model with stochastic vector $\hat{\mathbf{a}}_i = (\hat{a}_{ij})_{j \in [n]} \in \mathbb{R}^n_+$
 - E.g., $\hat{\bf a}_i=(1/2,1/3,1/6)\,\Rightarrow$ man 1 is 3 times more attractive to woman i (ex ante) than man 2
 - $\hat{\mathbf{A}}$: matrix with $\hat{\mathbf{a}}_i$'s as rows; similarly $\hat{\mathbf{B}}$ for men
- Stability: no blocking pairs
- Rank of an agent: index of her matched partner in her preference

A B M A B M



- *n* agents on each side ("women" and "men") with **logit-based** random preferences
 - Woman *i*'s preference is generated from a logit model with stochastic vector $\hat{\mathbf{a}}_i = (\hat{a}_{ij})_{j \in [n]} \in \mathbb{R}^n_+$
 - E.g., $\hat{\bf a}_i=(1/2,1/3,1/6)\,\Rightarrow$ man 1 is 3 times more attractive to woman i (ex ante) than man 2
 - $\hat{\mathbf{A}}$: matrix with $\hat{\mathbf{a}}_i$'s as rows; similarly $\hat{\mathbf{B}}$ for men
- Stability: no blocking pairs
- Rank of an agent: index of her matched partner in her preference
- Working examples:



- *n* agents on each side ("women" and "men") with **logit-based** random preferences
 - Woman *i*'s preference is generated from a logit model with stochastic vector $\hat{\mathbf{a}}_i = (\hat{a}_{ij})_{j \in [n]} \in \mathbb{R}^n_+$
 - E.g., $\hat{\bf a}_i=(1/2,1/3,1/6)$ \Rightarrow man 1 is 3 times more attractive to woman i (ex ante) than man 2
 - $\hat{\mathbf{A}}$: matrix with $\hat{\mathbf{a}}_i$'s as rows; similarly $\hat{\mathbf{B}}$ for men
- Stability: no blocking pairs
- Rank of an agent: index of her matched partner in her preference
- Working examples:
 - Homogeneous case (uniform preferences): $\hat{\mathbf{A}} = \hat{\mathbf{B}} = \frac{1}{n} \mathbf{1}_n$

• • = • • = •



- *n* agents on each side ("women" and "men") with **logit-based** random preferences
 - Woman *i*'s preference is generated from a logit model with stochastic vector $\hat{\mathbf{a}}_i = (\hat{a}_{ij})_{j \in [n]} \in \mathbb{R}^n_+$
 - E.g., $\hat{\bf a}_i=(1/2,1/3,1/6)$ \Rightarrow man 1 is 3 times more attractive to woman i (ex ante) than man 2
 - $\hat{\mathbf{A}}$: matrix with $\hat{\mathbf{a}}_i$'s as rows; similarly $\hat{\mathbf{B}}$ for men
- Stability: no blocking pairs
- Rank of an agent: index of her matched partner in her preference
- Working examples:
 - Homogeneous case (uniform preferences): $\hat{\mathbf{A}} = \hat{\mathbf{B}} = \frac{1}{n} \mathbf{1}_n$
 - Community structures: $\mathbf{A} = \mathbf{B} = \frac{2}{n} \begin{pmatrix} \mathbf{1}_{n/2} & \mathbf{0}_{n/2} \\ \mathbf{0}_{n/2} & \mathbf{1}_{n/2} \end{pmatrix}$

A B M A B M



- *n* agents on each side ("women" and "men") with **logit-based** random preferences
 - Woman i's preference is generated from a logit model with stochastic vector $\hat{\mathbf{a}}_i = (\hat{a}_{ij})_{j \in [n]} \in \mathbb{R}^n_+$
 - E.g., $\hat{\bf a}_i=(1/2,1/3,1/6)$ \Rightarrow man 1 is 3 times more attractive to woman i (ex ante) than man 2
 - $\hat{\mathbf{A}}$: matrix with $\hat{\mathbf{a}}_i$'s as rows; similarly $\hat{\mathbf{B}}$ for men
- Stability: no blocking pairs
- Rank of an agent: index of her matched partner in her preference
- Working examples:
 - Homogeneous case (uniform preferences): $\hat{\mathbf{A}} = \hat{\mathbf{B}} = \frac{1}{n} \mathbf{1}_n$
 - Community structures: $\mathbf{A} = \mathbf{B} = \frac{2}{n} \begin{pmatrix} \mathbf{1}_{n/2} & \mathbf{0}_{n/2} \\ \mathbf{0}_{n/2} & \mathbf{1}_{n/2} \end{pmatrix}$
 - Public popularities: $\hat{\mathbf{a}}_i = \mathbf{a}$ for all i, $\hat{\mathbf{b}}_j = \mathbf{b}$ for all j

Understand the welfare structure in this general model:

イロト イヨト イヨト イヨ

Understand the welfare structure in this general model:

- Which agents are more competitive/popular?
- Is the market well connected?
- What do rank distributions look like?
- What welfare should we expect on each side?

Understand the welfare structure in this general model:

- Which agents are more competitive/popular?
- Is the market well connected?
- What do rank distributions look like?
- What welfare should we expect on each side?

What do those even mean in a heterogeneous market?

Understand the welfare structure in this general model:

- Which agents are more competitive/popular?
- Is the market well connected?
- What do rank distributions look like?
- What welfare should we expect on each side?

What do those even mean in a heterogeneous market?

Goal: Under certain **connectivity condition** on the market, all stable outcomes behave in a **certain way** that reflects **fitness of agents**.

Understand the welfare structure in this general model:

- Which agents are more competitive/popular?
- Is the market well connected?
- What do rank distributions look like?
- What welfare should we expect on each side?

What do those even mean in a heterogeneous market?

Goal: Under certain **connectivity condition** on the market, all stable outcomes behave in a **certain way** that reflects **fitness of agents**.

> Ultimately helping market designers to better analyze and enhance market efficiency, e.g., revealing inefficiency, identifying disadvantaged groups, etc.

Contents

Motivation: key questions

- 2 Preparation: fitness of agents and contiguity of market
- 3 Results: characterizing welfare distribution

4 Sketch of analysis

5 Open directions

Literature

• Uniformly random preferences

- Number of stable matchings ($\approx \frac{n \log n}{e}$), optimal and pessimal average ranks of each side (Pittel 1989)
- Number of stable partners; "law of hyperbola": product of average ranks of the two sides $\approx n$ (Pittel 1992)
- Unbalanced markets: short side advantage (Ashlagi, Kanoria, and Leshno 2017; Cai and Thomas 2022)
- Markets with public scores (Immorlica and Mahdian 2015; Kojima and Pathak 2009; Ashlagi, Braverman, and Hassidim 2014,etc.)
- Distribution of match characteristics (Menzel 2015; Pęski 2017)
 - Probability of a pair being matched
 - General preference model, many agents of each type

A B M A B M

Preparation: fitness of agents and contiguity of market

Ashlagi, Braverman, Zhao

Welfare Structure in Two-sided Random Matching Markets

7 / 20

→ < Ξ → </p>

Fitness of agents

Question

Can we characterize ex-ante competitiveness ("average popularity") of agents in a market **independent of** the realized stable matching?

• • = • • =

Fitness of agents

Question

Can we characterize ex-ante competitiveness ("average popularity") of agents in a market **independent of** the realized stable matching?

Example (Public fitness)											
$\hat{\mathbf{A}} = egin{array}{c} Amy \ Betty \ Cindy \end{array}$	${\rm Dan} \\ \begin{pmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{pmatrix}$	Evan 1/3 1/3 1/3	$\begin{array}{c} {\rm Fran} \\ 1/6 \\ 1/6 \\ 1/6 \end{array} \right)$	$\hat{\mathbf{B}}=rac{ extsf{Dan}}{ extsf{Fran}}$	$\begin{array}{c} \text{Amy} \\ \left(\begin{array}{c} 1/5 \\ 1/5 \\ 1/5 \end{array} \right) \end{array}$	Betty 3/10 3/10 3/10	$\left.\begin{array}{c} \texttt{Cindy}\\ 1/2\\ 1/2\\ 1/2\end{array}\right)$				

→ < Ξ → </p>

Fitness of agents

Question

Can we characterize ex-ante competitiveness ("average popularity") of agents in a market **independent of** the realized stable matching?

Example (Public fitness)											
$\hat{\mathbf{A}} = egin{array}{c} \mathtt{Amy} \ \mathtt{Betty} \ \mathtt{Cindy} \end{array}$	Dan $\begin{pmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{pmatrix}$	Evan 1/3 1/3 1/3	$\begin{array}{c} {\rm Fram} \\ 1/6 \\ 1/6 \\ 1/6 \end{array} \right)$	$\hat{\mathbf{B}}=rac{ extsf{Dan}}{ extsf{Fran}}$	Amy $\begin{pmatrix} 1/5 \\ 1/5 \\ 1/5 \\ 1/5 \end{pmatrix}$	Betty 3/10 3/10 3/10	$\begin{array}{c} {\rm Cindy} \\ 1/2 \\ 1/2 \\ 1/2 \end{array} \right)$				

What about markets with non-public fitness?

Fitness through mutual scaling

Observation

Rescaling the rows of $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$ (i.e., $\mathbf{A} \leftarrow \operatorname{diag}(\phi)\hat{\mathbf{A}}$ and $\mathbf{B} \leftarrow \operatorname{diag}(\psi)\hat{\mathbf{B}}$ with $\phi, \psi \in \mathbb{R}^n_+$) has no impact on the preference model.

Fitness through mutual scaling

Observation

Rescaling the rows of $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$ (i.e., $\mathbf{A} \leftarrow \operatorname{diag}(\phi)\hat{\mathbf{A}}$ and $\mathbf{B} \leftarrow \operatorname{diag}(\psi)\hat{\mathbf{B}}$ with $\phi, \psi \in \mathbb{R}^n_+$) has no impact on the preference model. There exist an **essentially unique** rescaling (ϕ, ψ) such that

 $\mathbf{M} = n\mathbf{A} \circ \mathbf{B}^{\top}$

is doubly stochastic (having unit row and column sums).

Fitness through mutual scaling

Observation

Rescaling the rows of $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$ (i.e., $\mathbf{A} \leftarrow \operatorname{diag}(\phi)\hat{\mathbf{A}}$ and $\mathbf{B} \leftarrow \operatorname{diag}(\psi)\hat{\mathbf{B}}$ with $\phi, \psi \in \mathbb{R}^n_+$) has no impact on the preference model. There exist an **essentially unique** rescaling (ϕ, ψ) such that

 $\mathbf{M} = n\mathbf{A} \circ \mathbf{B}^{\top}$

is doubly stochastic (having unit row and column sums).

• M: mutual matrix

• ϕ and ψ : (anti-)fitness of the women and men

Example: public popularity (revisited)



10 / 20

Example: public popularity (revisited)

$$\hat{\mathbf{A}} = \begin{array}{cccc} & \text{Dan Evan Fran} & \text{Amy Betty Cindy} \\ \hat{\mathbf{A}} = \begin{array}{cccc} & \text{Amy} & \left(\begin{array}{cccc} 1/2 & 1/3 & 1/6 \\ 1/2 & 1/3 & 1/6 \\ 1/2 & 1/3 & 1/6 \end{array} \right) & \hat{\mathbf{B}} = \begin{array}{cccc} & \text{Dan} & \left(\begin{array}{cccc} 1/5 & 3/10 & 1/2 \\ 1/5 & 3/10 & 1/2 \\ 1/5 & 3/10 & 1/2 \end{array} \right) \\ \phi \propto \begin{array}{cccc} & \text{Amy} & \left(\begin{array}{cccc} 1/2 \\ 1/2 & 1/3 & 1/6 \end{array} \right) & \hat{\mathbf{B}} = \begin{array}{cccc} & \text{Evan} & \left(\begin{array}{cccc} 1/5 & 3/10 & 1/2 \\ 1/5 & 3/10 & 1/2 \end{array} \right) \\ \phi \propto \begin{array}{cccc} & \text{Amy} & \left(\begin{array}{cccc} 1/2 \\ 1/2 & 1/3 & 1/6 \end{array} \right) & \hat{\mathbf{B}} = \begin{array}{ccccc} & \text{Evan} & \left(\begin{array}{cccc} 2 \\ 3 \\ 6 \end{array} \right) \\ \psi \propto \begin{array}{cccc} & \text{Evan} & \left(\begin{array}{ccccc} 2 \\ 3 \\ 6 \end{array} \right) \end{array}$$

Ashlagi, Braverman, Zhao

Welfare Structure in Two-sided Random Matching Markets

10 / 20

2

▲□ ▶ ▲ □ ▶ ▲ □ ▶

Example: public popularity (revisited)

$$\hat{\mathbf{A}} = \begin{array}{cccc} & \text{Dan Evan Fran} & \text{Amy Betty Cindy} \\ \hat{\mathbf{A}} = \begin{array}{cccc} \text{Amy} & \begin{pmatrix} 1/2 & 1/3 & 1/6 \\ 1/2 & 1/3 & 1/6 \\ 1/2 & 1/3 & 1/6 \\ \end{pmatrix} & \hat{\mathbf{B}} = \begin{array}{cccc} \text{Dan} & \begin{pmatrix} 1/5 & 3/10 & 1/2 \\ 1/5 & 3/10 & 1/2 \\ 1/5 & 3/10 & 1/2 \\ 1/5 & 3/10 & 1/2 \\ \end{pmatrix} \\ \phi \propto \begin{array}{cccc} \text{Amy} & \begin{pmatrix} 1/2 \\ 1/3 \\ 1/5 \\ \end{array} \\ \phi \propto \begin{array}{cccc} \text{Betty} \\ 1/3 \\ 1/5 \\ \end{array} \\ \begin{pmatrix} 1/2 \\ 1/3 \\ 1/5 \\ \end{array} \\ \psi \propto \begin{array}{cccc} \text{Evan} \\ \text{Evan} \\ \text{Fran} \\ \begin{pmatrix} 2 \\ 3 \\ 6 \\ \end{array} \\ \end{pmatrix} \end{array}$$

Ashlagi, Braverman, Zhao

Welfare Structure in Two-sided Random Matching Markets

10 / 20

2

▲□ ▶ ▲ □ ▶ ▲ □ ▶

Connectivity of the market

Example (Sub-markets)

Consider
$$\mathbf{A} = \mathbf{B} = \frac{2}{n} \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{pmatrix}$$
. The market decomposes into two independent sub-markets

11 / 20

* (四) * * 注 * * 注

Connectivity of the market

Example (Sub-markets)

Consider
$$\mathbf{A} = \mathbf{B} = \frac{2}{n} \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{pmatrix}$$
. The market decomposes into two independent sub-markets.

We can only hope to characterize markets that are "well connected".

11 / 20

通 ト イ ヨ ト イ ヨ ト
Connectivity of the market

Example (Sub-markets)

Consider
$$\mathbf{A} = \mathbf{B} = \frac{2}{n} \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{pmatrix}$$
. The market decomposes into two independent sub-markets

We can only hope to characterize markets that are "well connected".

Assumption (Connectivity)

There exists $C < \infty$ independent of n such that

$$\frac{a_{ij}}{a_{ij'}}, \frac{b_{ji}}{b_{ji'}} \le C \qquad \forall i, i', j, j'.$$

• The uniform case is the special case when ${\cal C}=1$

Connectivity of the market

Example (Sub-markets)

Consider
$$\mathbf{A} = \mathbf{B} = \frac{2}{n} \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{pmatrix}$$
. The market decomposes into two independent sub-markets

We can only hope to characterize markets that are "well connected".

Assumption (Connectivity)

There exists $C < \infty$ independent of n such that

$$\frac{a_{ij}}{a_{ij'}}, \frac{b_{ji}}{b_{ji'}} \le C \qquad \forall i, i', j, j'.$$

- The uniform case is the special case when ${\cal C}=1$
- \bullet Bounded spectral gaps of $\mathbf{A}, \mathbf{B},$ and \mathbf{M} are probably sufficient

Results: characterizing welfare distribution

Ashlagi, Braverman, Zhao

Welfare Structure in Two-sided Random Matching Markets

12 / 20

Welfare characterization

We measure satisfaction of an agent by her rank-to-(anti-)fitness (RTF) ratio:

- Given an agent's fitness, smaller rank means happier
- Given an agent's rank, larger (anti-)fitness means happier

Welfare characterization

We measure satisfaction of an agent by her rank-to-(anti-)fitness (RTF) ratio:

- Given an agent's fitness, smaller rank means happier
- Given an agent's rank, larger (anti-)fitness means happier

Theorem (Informal)

Whp, in every stable matching, the followings hold:

Product of two sides' average RTFs is close to n;
 Trade-off between the two sides

Welfare characterization

We measure satisfaction of an agent by her rank-to-(anti-)fitness (RTF) ratio:

- Given an agent's fitness, smaller rank means happier
- Given an agent's rank, larger (anti-)fitness means happier

Theorem (Informal)

Whp, in every stable matching, the followings hold:

- Product of two sides' average RTFs is close to n;
 Trade-off between the two sides
- Empirical distribution of RTFs on each side is close to exponential.
 - Anti-concentration due to stability constraint

Each agent's preference is a uniformly random ordering of the opposite side.

Corollary (Informal, uniform case)

Whp, in every stable matching, the followings hold:

- **9** Product of two sides' average **ranks** is close to n; (law of hyperbola, Pittel 1992)
- Empirical distribution of ranks on each side is close to exponential.

.

Each agent's preference is a uniformly random ordering of the opposite side.

Corollary (Informal, uniform case)

Whp, in every stable matching, the followings hold:

- **9** Product of two sides' average **ranks** is close to n; (law of hyperbola, Pittel 1992)
 - Empirical distribution of ranks on each side is close to exponential.

Simulation: $n = 10^4$



Each agent's preference is a uniformly random ordering of the opposite side.

Corollary (Informal, uniform case)

Whp, in every stable matching, the followings hold:

- **9** Product of two sides' average **ranks** is close to n; (law of hyperbola, Pittel 1992)
 - Empirical distribution of ranks on each side is close to exponential.

Simulation: $n = 10^4$



Each agent's preference is a uniformly random ordering of the opposite side.

Corollary (Informal, uniform case)

Whp, in every stable matching, the followings hold:

9 Product of two sides' average **ranks** is close to n; (law of hyperbola, Pittel 1992)

14 / 20



Each agent's preference is a uniformly random ordering of the opposite side.

Corollary (Informal, uniform case)

Whp, in every stable matching, the followings hold:

9 Product of two sides' average **ranks** is close to n; (law of hyperbola, Pittel 1992)

14 / 20



Each agent's preference is a uniformly random ordering of the opposite side.

Corollary (Informal, uniform case)

Whp, in every stable matching, the followings hold:

- **9** Product of two sides' average **ranks** is close to n; (law of hyperbola, Pittel 1992)
 - Empirical distribution of ranks on each side is close to exponential.



Each agent's preference is a uniformly random ordering of the opposite side.

Corollary (Informal, uniform case)

Whp, in every stable matching, the followings hold:

9 Product of two sides' average **ranks** is close to n; (law of hyperbola, Pittel 1992)

14 / 20



Each agent's preference is a uniformly random ordering of the opposite side.

Corollary (Informal, uniform case)

Whp, in every stable matching, the followings hold:

9 Product of two sides' average **ranks** is close to n; (law of hyperbola, Pittel 1992)

14 / 20



Each agent's preference is a uniformly random ordering of the opposite side.

Corollary (Informal, uniform case)

Whp, in every stable matching, the followings hold:

9 Product of two sides' average **ranks** is close to n; (law of hyperbola, Pittel 1992)

14 / 20



Each agent's preference is a uniformly random ordering of the opposite side.

Corollary (Informal, uniform case)

Whp, in every stable matching, the followings hold:

- **O** Product of two sides' average **ranks** is close to *n*; (law of hyperbola, Pittel 1992)
 - Empirical distribution of **ranks** on each side is close to exponential.





• The advantaged side has the **same average rank** in both matchings (coupon collector), but very **different distributions**

15 / 20



- The advantaged side has the **same average rank** in both matchings (coupon collector), but very **different distributions**
- The deferred acceptance mechanism is **symmetric**: ordering of proposing does not matter; RSD is not: ordering matters

15 / 20



- The advantaged side has the **same average rank** in both matchings (coupon collector), but very **different distributions**
- The deferred acceptance mechanism is **symmetric**: ordering of proposing does not matter; RSD is not: ordering matters

15 / 20



- The advantaged side has the **same average rank** in both matchings (coupon collector), but very **different distributions**
- The deferred acceptance mechanism is symmetric: ordering of proposing does not matter; RSD is not: ordering matters
 Max entropy heuristics: in stable matchings, average rank captures "all info"; in RSD, there is extra info



- The advantaged side has the **same average rank** in both matchings (coupon collector), but very **different distributions**
- The deferred acceptance mechanism is symmetric: ordering of proposing does not matter; RSD is not: ordering matters
 Max entropy heuristics: in stable matchings, average rank captures "all info"; in RSD, there is extra info
- Average rank on the disadvantaged side is sublinear in WOSM, yet linear in RSD

Sketch of analysis

Ashlagi, Braverman, Zhao

Welfare Structure in Two-sided Random Matching Markets

16 / 20

æ

イロト イヨト イヨト イヨ

Rank is discrete – hard to work with, use a continuous proxy:

• Each woman i generates a value $X_{ij} \sim \mathrm{Exp}(na_{ij})$ independently for each man j

A = A A =
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A

Rank is discrete – hard to work with, use a continuous proxy:

- Each woman i generates a value $X_{ij} \sim \mathrm{Exp}(na_{ij})$ independently for each man j
- Woman i prefers man j to $j' \iff X_{ij} < X_{ij'}$

• • = • • = •

Rank is discrete – hard to work with, use a continuous proxy:

- Each woman i generates a value $X_{ij} \sim \mathrm{Exp}(na_{ij})$ independently for each man j
- Woman i prefers man j to $j' \Longleftrightarrow X_{ij} < X_{ij'}$
- Men's preferences are generated analogously

A = A = A

Rank is discrete – hard to work with, use a continuous proxy:

- Each woman i generates a value $X_{ij} \sim \mathrm{Exp}(na_{ij})$ independently for each man j
- Woman i prefers man j to $j' \Longleftrightarrow X_{ij} < X_{ij'}$
- Men's preferences are generated analogously

A = A = A

Rank is discrete – hard to work with, use a continuous proxy:

- Each woman i generates a value $X_{ij} \sim \mathrm{Exp}(na_{ij})$ independently for each man j
- Woman i prefers man j to $j' \Longleftrightarrow X_{ij} < X_{ij'}$
- Men's preferences are generated analogously

 \succ Smaller value and smaller (anti-)fitness \iff smaller rank. Suffices to consider empirical distribution of values

• • = • • = •

• Consider $\{\texttt{Amy} \leftrightarrow \texttt{Dan}, \texttt{Betty} \leftrightarrow \texttt{Evan}, \texttt{Cindy} \leftrightarrow \texttt{Fran}\}$

18 / 20

→ ∃ → < ∃</p>

- Consider {Amy \leftrightarrow Dan, Betty \leftrightarrow Evan, Cindy \leftrightarrow Fran}
- Let x_1, \ldots, x_3 and y_1, \ldots, y_3 be the realized values

- Consider {Amy \leftrightarrow Dan, Betty \leftrightarrow Evan, Cindy \leftrightarrow Fran}
- Let x_1,\ldots,x_3 and y_1,\ldots,y_3 be the realized values
- Likelihood of stability given values

$$\begin{aligned} (1 - \mathbb{P}(\texttt{Amy},\texttt{Evan block}))(1 - \mathbb{P}(\texttt{Amy},\texttt{Fran block})) \cdots (1 - \mathbb{P}(\texttt{Cindy},\texttt{Evan block})) \\ &= (1 - F_{na_{12}}(x_1)F_{nb_{21}}(y_2))(1 - F_{na_{13}}(x_1)F_{nb_{31}}(y_3)) \cdots (1 - F_{na_{32}}(x_3)F_{nb_{23}}(y_2)) \\ &\approx \prod_{i \neq j} (1 - n^2 a_{ij}b_{ji}x_iy_j) \approx \prod_{i,j} (1 - nm_{ij}x_iy_j) \approx \exp(-n\mathbf{x}^\top \mathbf{My}) \end{aligned}$$

where $F_{\lambda}(z) = 1 - e^{-\lambda z}$ denotes exponential CDF

A B M A B M

- Consider {Amy \leftrightarrow Dan, Betty \leftrightarrow Evan, Cindy \leftrightarrow Fran}
- Let x_1,\ldots,x_3 and y_1,\ldots,y_3 be the realized values
- Likelihood of stability given values

$$\begin{aligned} (1 - \mathbb{P}(\texttt{Amy},\texttt{Evan block}))(1 - \mathbb{P}(\texttt{Amy},\texttt{Fran block})) \cdots (1 - \mathbb{P}(\texttt{Cindy},\texttt{Evan block})) \\ &= (1 - F_{na_{12}}(x_1)F_{nb_{21}}(y_2))(1 - F_{na_{13}}(x_1)F_{nb_{31}}(y_3)) \cdots (1 - F_{na_{32}}(x_3)F_{nb_{23}}(y_2)) \\ &\approx \prod_{i \neq j} (1 - n^2 a_{ij}b_{ji}x_iy_j) \approx \prod_{i,j} (1 - nm_{ij}x_iy_j) \approx \exp(-n\mathbf{x}^{\top}\mathbf{My}) \approx \exp(-\sum_{i \neq j} x_iy_j) \end{aligned}$$

where $F_{\lambda}(z) = 1 - e^{-\lambda z}$ denotes exponential CDF

Use contiguity assumption to show nx[⊤]My ≈ ∑_{i,j} x_iy_j for likely matchings
 → Happens when x, y are both in the principal eigenspace of M

A B M A B M

- Consider {Amy \leftrightarrow Dan, Betty \leftrightarrow Evan, Cindy \leftrightarrow Fran}
- Let x_1,\ldots,x_3 and y_1,\ldots,y_3 be the realized values
- Likelihood of stability given values

$$\begin{aligned} (1 - \mathbb{P}(\texttt{Amy},\texttt{Evan block}))(1 - \mathbb{P}(\texttt{Amy},\texttt{Fran block})) \cdots (1 - \mathbb{P}(\texttt{Cindy},\texttt{Evan block})) \\ &= (1 - F_{na_{12}}(x_1)F_{nb_{21}}(y_2))(1 - F_{na_{13}}(x_1)F_{nb_{31}}(y_3)) \cdots (1 - F_{na_{32}}(x_3)F_{nb_{23}}(y_2)) \\ &\approx \prod_{i \neq j} (1 - n^2 a_{ij}b_{ji}x_iy_j) \approx \prod_{i,j} (1 - nm_{ij}x_iy_j) \approx \exp(-n\mathbf{x}^{\top}\mathbf{My}) \approx \exp(-\sum_{i \neq j} x_iy_j) \end{aligned}$$

where $F_{\lambda}(z) = 1 - e^{-\lambda z}$ denotes exponential CDF

- Use contiguity assumption to show nx[⊤]My ≈ ∑_{i,j} x_iy_j for likely matchings
 ≻ Happens when x, y are both in the principal eigenspace of M
- In posterior, **x** conditional upon stability "looks like" i.i.d. $\operatorname{Exp}(\sum y_j)$ samples

• • = • • = •

- Consider {Amy \leftrightarrow Dan, Betty \leftrightarrow Evan, Cindy \leftrightarrow Fran}
- Let x_1,\ldots,x_3 and y_1,\ldots,y_3 be the realized values
- Likelihood of stability given values

$$\begin{aligned} (1 - \mathbb{P}(\texttt{Amy},\texttt{Evan block}))(1 - \mathbb{P}(\texttt{Amy},\texttt{Fran block})) \cdots (1 - \mathbb{P}(\texttt{Cindy},\texttt{Evan block})) \\ &= (1 - F_{na_{12}}(x_1)F_{nb_{21}}(y_2))(1 - F_{na_{13}}(x_1)F_{nb_{31}}(y_3)) \cdots (1 - F_{na_{32}}(x_3)F_{nb_{23}}(y_2)) \\ &\approx \prod_{i \neq j} (1 - n^2 a_{ij}b_{ji}x_iy_j) \approx \prod_{i,j} (1 - nm_{ij}x_iy_j) \approx \exp(-n\mathbf{x}^\top \mathbf{My}) \approx \exp(-\sum_{i \neq j} x_iy_j) \end{aligned}$$

where $F_{\lambda}(z) = 1 - e^{-\lambda z}$ denotes exponential CDF

- Use contiguity assumption to show $n\mathbf{x}^{\top}\mathbf{M}\mathbf{y} \approx \sum_{i,j} x_i y_j$ for likely matchings > Happens when \mathbf{x}, \mathbf{y} are both in the principal eigenspace of \mathbf{M}
- In posterior, **x** conditional upon stability "looks like" i.i.d. $\operatorname{Exp}(\sum y_j)$ samples
- Use standard concentration inequalities and union bound to finish proof

Discussion

• Summary: Characterization of rank/welfare distribution in stable matchings

- Global trade-off between the sides
- Intrinsic quality of agents
- Exponential histogram of RTF

• • = • • = •

Discussion

- Summary: Characterization of rank/welfare distribution in stable matchings
 - Global trade-off between the sides
 - Intrinsic quality of agents
 - Exponential histogram of RTF
- Results and analysis extend to almost stable matchings, including almost balanced markets (with sublinear imbalance)

4 3 > 4 3

Discussion

- Summary: Characterization of rank/welfare distribution in stable matchings
 - Global trade-off between the sides
 - Intrinsic quality of agents
 - Exponential histogram of RTF
- Results and analysis extend to almost stable matchings, including almost balanced markets (with sublinear imbalance)
 Look at large sub-markets
- The connectivity condition can be relaxed
Open directions

- Many-to-one and many-to-many matchings, multi-sided matchings, correlated preferences, etc.
- Generalization of the connectivity condition (analogous to expansion of graphs)
- Empirical evidence (e.g., from NRMP)
- Efficient algorithms for inferring/learning fitness and connectivity from *ex-post* observations (preferences and outcomes)

• • = • • = •

Open directions

- Many-to-one and many-to-many matchings, multi-sided matchings, correlated preferences, etc.
- Generalization of the connectivity condition (analogous to expansion of graphs)
- Empirical evidence (e.g., from NRMP)
- Efficient algorithms for inferring/learning fitness and connectivity from *ex-post* observations (preferences and outcomes)

Goal: Help identify market failures and enhance market efficiency

A B b A B b

Open directions

- Many-to-one and many-to-many matchings, multi-sided matchings, correlated preferences, etc.
- Generalization of the connectivity condition (analogous to expansion of graphs)
- Empirical evidence (e.g., from NRMP)
- Efficient algorithms for inferring/learning fitness and connectivity from *ex-post* observations (preferences and outcomes)

Goal: Help identify market failures and enhance market efficiency

Thank you!

A B b A B b